

Fig 1a: Pass@1 Rate by Agent  
(Averaged Across Models)

